Some Nonobvious Consequences of Monitoring Time in Conversation[1]

Stanley Feldstein
University of Maryland Baltimore County

...time is the medium of narration, as it is the medium of life.
Thomas Mann

...perhaps the most important function of time...is as a matrix for embedding the behavior stream.
Eugene Galanter

This chapter is ultimately about mutual influence in communication. But first, I want to describe the very simple behavior that reflects such influence, some of the consequences of examining this behavior, and some of its theoretical implications, although not necessarily in that order. The behavior is the way in which we parse time in conversational interactions. By that, I mean the length of time taken by our pauses, our bursts of speech, the silences that usually intervene when one of us stops talking and the other starts, and by our intrusions when it's not our turn. It is not obvious to most people that simply the time taken by our pauses and speech bursts and the rest of the elements that characterize the stream of speech carries a surprising amount of information about us. It is not the kind of behavior to which most of us pay much attention unless we are actors or we meet a person who talks too much or too little. But it is the very fact that we are usually unaware of the behavior that makes it so useful for the investigation of psychological states.

Think about what speech would sound like if it were simply a sequence of sounds and silences. If we were to listen and pay attention to a sentence spoken in a language that we do not understand, it would, ordinarily, seem as if it was a sequence of brief silences and of sounds that varied in duration, pitch, and loudness. If we were to strip the sequence of its pitch, intonation countour, and volume but maintain the durations of the sounds and silences, that is, control them by transforming them into sounds of a single frequency and a particular decibel level, we would hear simply a sequence of

[1]Most of the research reported in this chapter that is concerned with the analyses of dialogues was done in collaboration with Drs. Joseph Jaffe, Beatrice Beebe (both of the Department of Communication Sciences, New York State Psychiatric Institute, Columbia University), Cynthia L. Crown (Department of Psychology, Xavier University, Ohio), and Michael D. Jasnow (George Washington University, Washington, D.C.), although they are not responsible for any reporting errors that I may have made. I am also grateful to Heather Quay and Terri Harold for their careful editing of [illegible]

tones and silences of varying durations. We call the tones, or sounds, *vocalizations* and the silences, *pauses*. That sequence represents the temporal pattern of the speech stream (Figure 1) and provides the basic data that my colleagues and I have spent the last several decades studying. After all those years, I still find it remarkable that such simple sequences are so informative.

Insert Figure 1 about here

It is important to know that this sound-silence sequence is derived automatically from speech by a specialized computer system[2] and is not generated by a listener with a stopwatch. Thus, within limits, it is not only a very precise and objective rendering of the sounds and silences in the speech stream but also, it does not require the constant alert attention of a listener.

Monologues

Of course, one aspect of the information of the sound-silence sequence that is less surprising than others is the fact that it mirrors speech rate. Although some of the evidence appears to be contradictory (*e.g.*, Ramsay, 1968; Scherer, 1979), rate of speech has been shown to be related to personality characteristics (Feldstein & Sloan, 1984; Ramsay, 1966), person perception (Miller, Maruyama, Beaber, & Valone, 1976), the Type-A coronary syndrome, (*e.g.*, Feldstein, Siegman, Simpson, Barkley, & Kobren, 1984) and certain affective disorders (Breznitz & Sherman, 1987). There are a number of ways to measure global rates of speech; one very common method is simply to count the number of words per minute. Allen, Anderson, and Hough (1968), for example, reported that approximately 150 words per minute was a "normal" speech rate. Other investigators have used syllables per minute or syllables per second as estimates of speech rate. Siegman (1987) used the silence quotient as a measure of speech rate, which is the sum of all pauses ≥ 2 sec. divided by the duration of the speech sample. Miller (1987) argued that the global speech rate is comprised of two components: the rate at which the speech sounds are produced (articulation rate) and the rate at which pauses are produced, or the numbers and durations of the pauses. However, it was

---

[2]The Automated Vocal Transaction Analyzer (AVTA, Cassotta, Feldstein, & Jaffe, 1964; Jaffe & Feldstein, 1970; Feldstein, BenDebba, & Alberti, 1974; Feldstein & Welkowitz, 1987) "listens" simultaneously to the speech of the two participants in a dialogue, performs an analog-to-digital conversion, and initially records, every quarter second, whether each participant is talking (coded as *1*) or silent (coded as *0*). The two resulting time series of *1*s and *0*s are then transformed into a sequence of four numbers that reflect the succession of the four observable dyadic states: both participants are silent (*0*), one participant is talking and the other is silent (*1*) or vice versa (*2*), both participants are talking (*3*). AVTA software converts these numbers into a set of vocal states and their descriptive statistics. For reliability estimates, see Feldstein & Welkowitz (1987).

Ramsey (1966, 1968) who measured the durations of the verbalizations and the silences among them to obtain a sound-silence sequence. He then used mean lengths of sounds and of silences as well as the ratio of the total sound time to the total silence time to distinguish, successfully, between reading aloud and speaking spontaneously, and between extraverts and introverts. In earlier work with speech rate (Feldstein, 1976), I compared, by means of multiple regression analyses, each of the following with global speech rate, *i.e.*, words per minute, of monologues: (a) the average sound duration, (b) the average silence duration , (c) the ratio of the average sound duration to the average silence duration, and (d) the ratio of the proportionality constant[3] of sounds to that of silences (PCR). The resulting coefficients were -.14, .31, .35, and .90, respectively.

The point is that a simple sequence of the sounds and silences that characterize the speech of one person contains sufficient information to yield a measure that is quite significantly related to global speech rate. Not only, incidentally, can the sequence be used to obtain a measure (the PCR) of speech rate, but persons who hear it can estimate, in terms of the adjectives *low, moderate*, and *fast*, the rate of the speech segment from which it was derived (Crown & Feldstein, 1991; Feldstein, Crown, & Jaffe, 1991).

The single sound-silence sequence also yields other types of information. It can reflect what might be called an individual's "state of consciousness." Krüger and his colleagues (Krüger, Rausche, Rimkus, & Voorath, 1991) , for example, built an instrument called the Logoport[4] which, like AVTA, extracts and codes the sound-silence sequence in speech. The sequences of individual speakers have been used by Krüger and his colleagues to determine the effects and course of drugs and alcohol on speech. A colleague and I (Feldstein & Weingartner, 1981) reviewed a selection of studies illuminating the effects of various kinds of drugs on the speech stream.

The single sound-silence sequence was used in studies of psychopathology. One study examined the behavior of five acutely psychotic patients (Glaister, Feldstein, & Pollack, 1980). The sequences were transformed into alternating series, computed by subtracting each silence from its preceding vocalization and summing the differences. The slopes of the resulting curves were found to be related to the interviewers' (a psychiatrist and psychologist) impressions of how ill the patients were and to the patients' impressions about their progress in the interviews. A number of other studies (Greden & Carroll, 1980; Teasdale, Fogarty, & Williams, 1980) have found that the pace of the speech sequence reflects serious depression and variations in the level of depression. Indeed, another study (Anderson, Jaffe, & Chen, 1992) uncovered a bimodal distribution of the speech rates (PCRs) of depressed patients, which

---

[3]A proportionality constant is the limiting ratio of the frequency of an event of duration *n*-1 to the frequency of event duration *n*, and it is identical to the probability of continuing the event in question.

[4]The Logoport was designed to track the sound-silence sequences of two interacting individuals and to combine the separate sequences into a single dialogue.

distinguished retarded from faster speech rates. It was further found that the patients with retarded rates were the ones who improved with drug therapy.

Dialogues

But all of this concerns the sound-silence sequence of a single individual. What information is contained in a pair of sequences, such as that in Figure 2, which represents two individuals talking? Inspection of the two sequences tells us only when one or the other of the partners is speaking, when they both speak simultaneously, and when they are mutually silent. A careful perusal of Figure 2 reveals that there are two kinds of simultaneous speech and two kinds of mutual silence and that neither of the two kinds can be formally distinguished from the other. However, the imposition of a simple rule clarifies the situation.

Insert Figure 2 about here

The rule codifies the necessary social convention which holds that in dialogues, individuals are supposed to take turns speaking and not speak at the same time. With this rule, we discover that there are emergent temporal characteristics that uniquely describe the dialogue, *dyadic* temporal characteristics which do not exist in a monologue. The rule is called the *turn* rule, and states that, in a dialogue, *the turn of each participant begins as soon as he or she vocalizes alone and ends as soon as the other participant vocalizes alone* (Jaffe & Feldstein, 1970). This rule preserves the individuality of each speaker while suggesting that it is the dyad that is the viable social unit, or "social molecule," (Schwab, 1960). Although a monologue can be derived from a dialogue, a dialogue cannot be derived from a monologue. That sequences are paired not randomly but are paired by virtue of a conversation has an immediate consequence. There are no longer simply pauses and vocalizations. Instead, there are *speaking turns*, vocalizations, simultaneous pauses, and simultaneous vocalizations (Figure 3).

Insert Figure 3 about here

Note that there are two types of simultaneous, or mutual, pauses. One type occurs within the turn of a single speaker, bounded by his or her vocalizations. It is an *intra*personal pause, if you will, but is simply called a *pause* (P). The other type is preceded by a vocalization of one speaker and terminated by the vocalization of the other speaker. It is an *inter*personal pause and is called a *switching pause* (SP), because it marks the end of one speaker's *turn* and the beginning of the other speaker's *turn*. There are also two types of simultaneous vocalizations, or speech. One type is a vocalization of the speaker who does *not* have the turn that begins and ends during a vocalization of the speaker who does and, therefore, it is not an interruption and does not result in a switch of speakers. It is called *noninterruptive simultaneous speech*

(NSS). The other type begins during a vocalization of the speaker who has the turn and ends after that speaker has stopped talking. It does, therefore, interrupt the speaker and, since the interrupter ends up talking alone, he or she obtains the turn. That type of simultaneous vocalization is called *interruptive simultaneous speech* (ISS).

The *speaking turn*, then, is a supraordinate behavior in that it permits the delineation of the other dyadic vocal states and includes them within its boundaries. It is an inherent property of a dialogue. Switching pauses, noninterruptive and interruptive simultaneous speech are also unique characteristics of a dialogue although they are assigned to the individual speakers. By virtue of the turn rule, switching pauses are assigned to the speaker whose turn ended, noninterruptive simultaneous speech is assigned to the speaker who does not have the turn, while interruptive simultaneous speech is assigned to the speaker whose turn it initiated. Pauses and vocalizations are, of course, assigned to the speaker within whose turn they occur.

Coordinated Interpersonal Timing

Given the turn rule, the two sequences yield five vocal states which describe the temporal structure or, if you will, the temporal rhythms[5] of a dialogue. But what information about the participants in the dialogue do these five states and, hence, the pair of sound-silence sequences, carry? It was early in our research that we discovered (Feldstein, Jaffe, & Cassotta, 1967; Cassotta, Feldstein, & Jaffe, 1967; also reported in Jaffe & Feldstein, 1970) what appeared to be the mutual influence of the participants on each other's time patterns. The phenomenon of mutual influence was not unknown (Homans, 1950; Newcomb, 1953) and had been tested experimentally in interview research relatively similar to our own studies of conversation (Matarazzo & Wiens, 1967). With the discovery of its occurrence in our own data, my colleagues and I along with a number of our students initiated a series of studies concerned with the relation of mutual influence or, as we called it then, *congruence*, to (a) certain personality characteristics such as psychological differentiation (Feldstein, Alberti, BenDebba, & Welkowitz, 1974), (b) sociometric indices of affection and wanting social contact (Crown, 1991; Marcus, Welkowitz, Feldstein, & Jaffe, 1970; Feldstein & Welkowitz, 1987), (c) vocal behavior (Natale, 1975; Welkowitz, Feldstein, Finkelstein,& Aylesworth, 1972), (d) territoriality (Martindale, 1971), (e) interpersonal perception (Crown, Feldstein,& Bond, 1982; Feldstein, 1982; Welkowitz & Feldstein, 1970), and (f) emotional relationships between participants (Crown, 1982). All of these studies have been reviewed elsewhere (Feldstein & Welkowitz, 1987; Crown & Feldstein, 1985). Other investigators also became interested in mutual influence and found that it occurred with other types of behavior, such as body movements in psychotherapy (e.g., Scheflen, 1964), gestures (LaFrance, 1979), speech rate (Webb, 1972), expressive behavior (Cappella & Greene, 1982), and even accents (Giles, Taylor, & Bourhis, 1973).

---

[5]Dr. Alex Heller of the Graduate Center of City University, NY, suggested that they be called the "heterochronic" rhythms of dialogue (personal communication).

The results of our own studies seemed to indicate that some degree of temporal congruence occurred in almost all conversations. Unfortunately, however, the statistical technique we used in many of the studies allowed us to talk only about congruence over many conversations. To determine the extent of congruence in a *single* conversation, we began to use time-series regression analysis (Gottman, 1981; Warner, 1992).[6] Moreover, as a consequence of both our own findings and those of others, we reformulated our thinking about the mutual influence revealed in dyadic time patterns. We renamed the phenomenon *coordinated interpersonal timing*, or CIT, which seemed to us a more explicit description. CIT addresses the relationship between the temporal patterns of two individuals who are engaged in a dialogue; *it exists when the patterns of one of them can be predicted from the patterns of the other.* We wondered whether there was a relation between CIT and cognition. We also wondered whether CIT began to occur in dyadic vocal interactions at a specified time, that is, whether an individual begins to engage in such coordinated timing at a particular age. Can we observe CIT in children? Are the prelinguistic sounds that infants make temporally coordinated with the vocal behavior of their mothers? If so, does that mean that CIT serves, perhaps, a more basic function than as a correlate of certain psychological states? Specifically, we began to wonder not only whether the extent to which CIT occurs is related not only to the nature of the participants' relationship, but also whether it is a necessary, possibly neurobiological, "built-in," aspect of social interaction. Does the capacity for CIT serve "...as a mechanism for the facilitation of social interaction" (Jasnow, Crown, Feldstein, Taylor, Beebe, & Jaffe, 1988, p. 355)? It seems like a reasonable conjecture. A review of the appropriate literature (Jasnow et al., 1988) reveals that, in invertebrates and simple vertebrates such as *drosophila*, crickets, toads, and the electric fish, both genetic and neurological factors play a role in the temporal organization of social behavior. Significantly, the temporal organization of these organisms is frequently critical to their survival! Links between the social organization and neural substrates of these organisms can often be tested directly through surgical excision and other means. Such direct tests cannot, for excellent reasons (ethical and otherwise), be performed with humans. We can try only indirectly to examine the linkages between human social interaction and their neurobiological substrates, a rarely conclusive but very profitable endeavor.

One such indirect approach is to study animals that engage in behaviors more directly relevant to our concern. Kirkpatrick and his colleagues (Kirkpatrick, Carter, Newman, & Insel, 1994), for example, have been attempting to delineate the neurobiology of gregariousness. They chose to investigate prairie voles as animals that display singularly monogamous relationships with their mates as part of their pattern of social behavior. Surgical lesions were made of the basolateral nucleus and the corticomedial amygdala, the latter of which had the effect of reducing the contact time

---

[6]A detailed discussion of our use of time-series regression analysis can be found in Feldstein, Jaffe, Beebe, Crown, Jasnow, Fox, & Gordon (1993).

of the treated voles with a familiar female and a conspecific pup. Thus, the results implicate the corticomedial amygdala in several components of the social behavior of prairie voles. A second experiment showed that more restricted lesions affected only the paternal behavior of the voles. This work is intriguing and suggestive. A host of other studies have investigated the link between neurobiology and temporal factors in the behaviors of various species across the phylogenetic range (see Jasnow et al., 1988).

Adult-Adult Conversations

The phenomenon of CIT was first recognized in studies of adult-adult interactions. In this context, we explored the influence of interpersonal relationships on CIT. A study by Crown (1991) hypothesized that the extent to which adult pairs engage in CIT depends at least in part on the nature of their attraction to each other or the lack of it. Three groups of university women participated in the study. One group consisted of pairs who liked each other, a second group, of pairs who disliked each other, and a third of pairs in which the partners were unacquainted with each other. These were not artificial groups. Unlike the majority of studies concerned with interpersonal attraction in which the pairs have met for the first time during the study (Berscheid, 1985), the pairs in this study already had a relationship history. The individuals were paired on the basis of a sociometric scale on which dormitory residents, before they knew about the study, were asked to identify other residents who they liked or disliked or didn't know. Even after the study was completed, there was very little change in the mutual feelings of the paired individuals. The aim of the study was to discover whether interpersonal attraction was related to CIT and specifically, whether liking or disliking one another influences the degree to which a pair engages in CIT in their conversations. The study not only analyzed the timing of vocal behavior, but also of visual, or gaze behavior.

The results show that the unacquainted group engaged in the greatest degree of vocal and visual temporal coordination. This group was followed in order by the "dislike" and the "like" groups. Initially, these results seemed counterintuitive; surely, the interactions of those pairs who liked each other should show the most CIT. We had explicitly, if casually, assumed that CIT was directly associated with interpersonal attraction in a relationship and with the functional adequacy of the relationship (Feldstein & Welkowitz, 1987). It was an assumption that also seemed to be implied in the work of others (*e.g.*,Chapple, 1970; Cohn & Tronick, 1988; Martin, 1981). The results of Crown's study provided the first hint of what, we shall see, the results of our infant studies ultimately made abundantly clear; the assumption is not tenable.

Crown interpreted her findings as evidence that CIT may serve a regulatory function, which comes into play when actual or perceived constraints upon the participants seem to warrant it. Crown asked her participants to converse for 30 minutes about any topics that interested them. At bottom, the individuals in each pair understood that they had to sustain a conversation for 30 minutes. The students who liked each other apparently had free and easy conversations which required relatively little external motivation and not very much structure. Those who disliked each other were perhaps more guarded and careful of what they were saying and how they were saying it. Undoubtedly, they found it harder to continue speaking for the full 30 minutes. The students who did not know each other had to explore topics of conversation and, even more than the pairs who disliked each other, must have made a concerted effort to be responsive and maintain the conversation for the required time. Although the results were unexpected, the carefulness of the "dislike" pairs and the effort of the unacquainted pairs do seem to explain the higher coefficients of CIT. But does the same configuration occur in children's conversations?

Preadolescent Conversations

In collaboration with Tiffany Field, who collected the data in Florida for another study, I explored the use of CIT in the interactions of preadolescent friends and acquaintances (Feldstein, 1991).[7] Two aims of the study are important here. The first was to discover whether CIT occurs in the dialogues of preadolescent pairs. If so, the second was to compare the CIT of friends and acquaintances. The few investigations that have examined the chronometry of children's conversations were concerned not only with whether the time patterns exhibited CIT. They were also concerned with whether the children's time patterns were influenced by age, gender, and ethnicity. One such study (Welkowitz, Bond, & Feldstein, 1984a) found that the vocal time patterns of Hawaiian children were stable indices of their conversational behavior, and that the patterns seem to vary as a function of the gender and ethnicity of the conversational pairs. Another study (Welkowitz, Bond, & Feldstein, 1984b) of Japanese-American children and adults in mixed- and same-gender pairs found that gender affected the time patterns of the adults but not of the children. Two other earlier studies (Garvey & BenDebba, 1974; Welkowitz, Cariffe, & Feldstein, 1976) seemed to indicate that the development of CIT is positively related to age. However, none of these studies involved preadolescents, and the techniques they used to assess CIT were relatively crude.

However, Welkowitz and her colleagues (Welkowitz, Bond, Feldman, & Tota, 1990) did publish a study of four-to-five-year olds interacting with each of their parents, and derived what they called *influence coefficients* by time-series regression analysis. The coefficients were calculated for the children and the parents for the durations of four vocal states: turns, pauses, switching pauses, and vocalizations. The switching pauses showed the greatest influence. The study included other psychological variables, but its primary aim was to examine degree and direction of mutual influence.

The preadolescent study estimated CIT by the use of time-series regression analysis. The participants were 30 female and 26 male preadolescents recruited by Field from two sixth-grade classes. Their ages averaged 11½ years and they had known each other for a little over four years. The selection of friends and acquaintances was

---

[7]Dr. Field is with the Mailman Center for Child Development of the University of Miami School of Medicine.

made on the basis of sociometric information obtained from the preadolescents. From among their classmates, they were asked who they knew "the best," and who they knew "the least." The pairs were assembled into three groups, two groups of same-gender pairs and one of mixed-gender pairs. Each engaged in a 10-minute, face-to-face interaction across a small table about any topic or topics that interested them. Each conversation was audiotaped and then analyzed in terms of the vocal states. The data were then subjected to multiple regression analyses.

No simple distinctions could be found between the groups of friends and acquaintances. Although the turn and switching-pause durations of the boys showed greater coordination with friends than with acquaintances, those of the girls showed the opposite. The same pattern characterized the boys' and girls' interruptive behavior (*i.e.*, ISS). Moreover, same-gender friends coordinated the durations of their interruptive behavior to a significantly greater extent than did mixed-gender friends. This is the only finding that indicates that interruptive behavior is sensitive to the gender composition of a pair, but only when mediated by the type of relationship between the partners. But apart from whether they were friends or acquaintances, or whether they were in mixed- or same-gender pairs, the girls' coordination of their pause durations was greater than that of the boys.

The groups of child pairs do not quite coincide with the adult groups in Crown's study. Strictly speaking, there was no group of "unacquainted" pairs among the children, and it is an assumption that the pairs of preadolescent "friends" liked each other. However, the girls in the acquainted pair group behaved like the unacquainted pairs of women in the Crown study, in that their coordination was greater than that of the pairs in the friends group. Perhaps if the Crown study had included men, they would have behaved like the boys in the Florida study.

Although unrelated to interpersonal coordination, a number of other intriguing findings came to light in the Florida study. One is that girls tended to take longer turns than boys, whether in mixed- or same-gender pairs. More interesting was the preadolescents' use of simultaneous speech. The boys engaged in more frequent ISS and NSS with acquaintances than with friends, whereas the girls engaged in more ISS and NSS with friends than with acquaintances. In general, friends used longer durations of ISS in their interactions than did acquaintances. However, female friends used longer ISS durations than did female acquaintances, and male friends used shorter ISS durations than did male acquaintances. Also, friends used longer ISS durations in same-gender pairs whereas acquaintances used longer ISS durations in mixed-gender pairs.

Interruptive behavior is important not only because it has been tagged as a significant predictor of coronary artery disease (Siegman, Feldstein, Tomasso, Ringel, & Lating, 1987), but also because it is thought to reflect dominance in interpersonal interactions. Much of the relevant literature (Deaux, 1976; Eakins & Eakins, 1978; Henley, 1977) suggests that men interrupt more frequently than do women, that women are interrupted more often than are men, and that, in mixed gender pairs, men interrupt women more often than women interrupt men. However Dindia's (1987), relatively brief but useful review of the literature, concluded that these experimental findings are based upon both inadequate designs and faulty statistical analyses. Her own study, which attempted to correct such problems, found that men do not interrupt more than women and that women do not get interrupted more than do men. She also found that although more interruptions occurred in mixed-gender pairs than in same-gender pairs, they were not related to the gender of the speaker.

These results suggest that the engagement of boys and girls in ISS depended upon whether they were friends or acquaintances. Even the ISS in mixed- or same-gender pairs appeared to differ as a function of whether the partners were friends or acquaintances. The finding that there were no differences simply as a function of gender accords with the findings of Dindia (1987) and is not at variance with those of Natale, Entin, & Jaffe, (1979). That there were no simple differences between the ISS of boys and girls argues against the seemingly prevalent notion that interruptive behavior reflects male dominance, although not against the possibility that the use of ISS may occasionally reflect interpersonal dominance.

Deaux (1977) proposed that the strategy of men in interactive contexts tends to be competitive whereas that of women tends to be affiliative. Could it be that boys are more competitive with acquaintances than with friends but girls are more affiliative with acquaintances and more competitive with friends? Or might it be that the use of ISS means different things to boys and girls?

Adult-Infant Interactions

The two studies I have just described seem to indicate that, in both adult and preadolescent dialogues, CIT occurs and varies lawfully with other important variables. But when does it start? Do prelinguistic infants engage in CIT with adults? With these questions in mind, we first determined whether infants coordinated their prelinguistic vocal behavior with the vocal behavior of their mothers and of strangers. We investigated (Jaffe, Feldstein, Beebe, Crown, & Jasnow, 1995) the play interactions of approximately 83 mothers and their 12-month old infants. We used very stringent exclusion criteria to make certain that the infants and their mothers were healthy and "normal." In addition to playing with their mothers, we also had the infants play with female strangers, allowing us to determine not only whether infants are able to engage in CIT not only with their mothers but also with women in general; and, further, if they are able to, whether the magnitude of their coordination is different for mothers and strangers. The mother-infant and stranger-infant pairs were seen in the infants' homes when they were six weeks, four months, and 12 months old. Their interactions took, on the average, about 11 minutes. When the infants were 12 months old, their interactions with their mothers and the strangers also included a puppet. On each occasion, after the infants interacted with their mothers and the strangers, the mothers and strangers conversed for about 20 minutes. Approximately half the total number of pairs were also seen in our New York laboratory.[8] All interactions were audiotaped and those in the laboratory were also videotaped. Neither the mothers nor the strangers knew about the variables the study examined.

The AVTA system automatically coded the recorded vocal interactions of the pairs, and then derived descriptive statistics for each of the vocal states. In addition to the descriptive statistics for the entire time of each interaction, we obtained five-second averages for each state. Thus, the participation of each individual in an interaction could be described in terms of six vectors of five-second averages. In a 10-minute interaction, each vector would have a sequence of 120 five-second averages (12 five-second units per minute) which would essentially describe that vocal behavior over the course of the 10 minutes. To find the extent to which a particular mother-infant pair coordinated the timing of a particular vocal behavior, such as pauses, we compared their pause vectors using a time-series regression analysis (TSR).[9] TSR does two things that are of special interest. It isolates and quantifies the extent to which an individual's immediately prior behavior influences his or her current behavior (called the autocorrelation). It also isolates and quantifies the extent to which an individual's behavior is influenced by (or coordinated with) that of the other participant (called the cross-correlation). In other words, it separates what might be considered "self influence" from "other influence." The TSR yields a squared semi-partial correlation coefficient that is an estimate of "other influence;" we call it the coefficient of CIT.

Six-week-old infants do not engage in much vocal activity, which made it unprofitable to examine the timing of their vocal behavior. Instead, we (Crown, Feldstein, Jasnow, Beebe, Wagman, Gordon, Fox, & Jaffe, unpublished manuscript) examined their interactions with their mother and with strangers and compared the timing of their gaze behavior (at six weeks) with the timing of the adults' vocal behavior. Forty-five of the total number of triads (mother, infant, and stranger) in our New York study were randomly selected for this study. The results indicated (a) that CIT occurs cross-modally not only between mothers and their infants but also between strangers and infants, and (b) that for several of the vocal states, the cross-modal coordination between the strangers and infants is greater than that between mothers and infants.

What emerges most clearly from the data is that by six weeks of age newborn infants already display a marked sensitivity to the temporal patterns encountered during the course of a social interaction. This ability to engage in CIT is more than imitative; the infants apparently extract the invariant temporal properties of the vocal modality in

---

[8]The laboratory is located in the Department of Communication Sciences of the New York State Psychiatric Institute, Columbia University, New York. Another laboratory of our group is the Interpersonal Communication Laboratory, located in the Department of Psychology at the University of Maryland Baltimore County.

[9]A detailed description of the TSR procedure is described in Feldstein, Jaffe, Beebe, Crown, Jasnow, Fox, & Gordon, 1993. An excellent discussion of the theory and technique of time-series analysis is provided by Warner (1988).

which their partners are operating and express the information in the visual modality. This finding supports the contention that human beings possess, from very early on, a temporal sensitivity (Lewkowicz, 1989) that plays an important role in social exchanges. The finding that is also consonant with a host of prior studies (Allan, Walker, Symonds, & Marcell, 1977; Bryant, Jones, Claxton & Perkins, 1972; Freides, 1974; Jones & Robinson, 1973; Rose, Gottfried, & Bridger, 1978; Turkewitz & McGuire, 1978) that demonstrate the ability of infants to engage in complex cross-modal integrations. More importantly, perhaps, is that this temporal coordination between an adult and a six-week-old infant is a *shared* rhythmic behavior. Many researchers have commented on the important role played by such phenomena in homo sapiens as well as in a myriad of species along the phylogenetic scale. Shaffer (1977), for example, observed that "What appears to matter most about 'successful' mother-infant interactions is above all the temporal integration of the two partners' responses and when we talk about mothers' sensitivity it is often this temporal characteristic that we have in mind" (p. 214). Our finding that temporal coordination occurs at as early an age as six weeks provides further support for the notion that CIT may be a biological mechanism. The anthropologist, Paul Byers (1976) wrote that "We can...imagine a human or animal world that is communicationally related through the sharing of time forms in multiple levels of behavioral organization" (p. 160). We suggest that the capacity to engage in CIT is a necessary (though not sufficient) condition for two human beings to enter into effective dialogue with each other, be they six weeks or 60 years old.

The findings of this study of gaze and vocal behavior also lend support to another implication of our work that should be made explicit. We chose to examine the timing of vocal states in an interaction not only because they play a critical role in human communication, but also because they comprise the only aspect of interactions that could, at the time we began our research, be easily and precisely tracked by a computer system. (With the sophisticated technology now available, it has become possible to begin automating the tracking and recording of the time course of gestures and other body movements.) We expect, however, that we would find the same relationships were we to examine the timing of other modalities involved in interactions, such as nonverbal behaviors and body movements.

By the age of four months, infants do engage in vocal interaction with adults, making it possible to examine the extent to which they and their adult partner coordinated their vocal behavior. We evaluated the significance of the extent of coordination in 12 groups of dyads: the mothers' coordination with their infants in the home, the mothers' coordination with their infants in the laboratory, the infants' coordination with their mothers in the home and laboratory, the strangers' coordination with the infants in the home and laboratory and vice versa, and the mothers' coordination with the strangers in the home and laboratory and vice versa. To accomplish these analyses, we considered each dyad a separate experiment and performed a meta analysis for each group. Meta analysis usually used to quantitatively summarize a set of separate experiments, all of which test the same directional hypothesis or hypotheses. For our data, we employed three types of meta analyses and tested each vocal state separately. The first meta analysis yielded a chi-square (χ^2) that assessed the degree of diversity among the coefficients of CIT in each group. The second yielded a *t* test that assessed the overall statistical significance of the coefficients in each group. The third yielded was a *t* test that determined whether the coefficients of a group were predominantly positive or negative, thereby indicating the direction of coordination.[10] These three types of meta analyses were performed for both the four- and 12-month data.

Because the results indicated that the average coefficient of CIT was not statistically significant for any of the 12 groups, we were forced to reconsider the possibility that CIT is not ubiquitous, even though the results of our studies of adult conversations seemed to demonstrate that most dyadic interactions exhibited CIT. In fact, the χ^2s yielded by the meta analyses revealed that the probabilities associated with the coefficients in each group varied enormously. These probabilities, which index the likelihood with which the coefficients could have occurred by chance, ranged from well below .05 to almost 1.00, meaning that although some of the coefficients probably did not occur by chance, others probably did. Our third meta analysis clearly showed that, for all the groups, pauses and vocalizations involved negative temporal coordination while switching pauses, ISS and NSS involve positive coordination. In other words, as the pauses and vocalizations of one speaker get longer, those of the other speaker get shorter. However, when the switching pauses and simultaneous speech segments of one speaker become longer (or shorter), those of the other speaker follow suit. Despite the lack of significant group averages, it should be noted that the coefficients are lawfully related to other important variables in the study (Jaffe at al., 1995). As we shall see, the coefficients derived from the four-month interactions predicted certain meaningful developmental outcomes of the infants when they were 12 months. But these were all group analyses. Examination of the separate interactions tells a somewhat different story.

For a particular pair of speakers, we can find no CIT or we can find one of two types of CIT: bidirectional coordination (in which case both participants coordinate the timing of their vocal behavior); or unidirectional coordination (in which one participant coordinates, while the other does not). If we look only at the adult-infant interactions when the infants were four months old, we find that 63% of them show bidirectional CIT. Approximately 28% show unidirectional coordination, 15% with the infants as the

---

[10] One type of meta analysis we used, for which we are indebted to Robert Rosenthal of Harvard University, involves three simple steps: (a) transforming the probability values associated with the coefficients of CIT to normal standard deviate scores, *i.e.*, *z* scores; (b) squaring each of the *z* scores to yield a chi-square with one degree of freedom; and (c) summing the chi-squares in each group to provide a chi-square with degrees of freedom equal to the number of chi-squares in the group. The other meta analysis was a *t* test for a single mean, for which beta coefficients produced by the TSRs were treated as scores. For this analysis, the absolute values of the original signed coefficients were used. The final type was also a *t* test for a single mean, but for this analysis the signed coefficients were used.

unilateral coordinators, and about 13% with the adults as the unilateral coordinators. Approximately 9% of the interactions showed no statistically significant levels of CIT.[11]

Analysis of the adult-infant interactions once the infants had reached the age of 12 months showed that only about 38% exhibited bidirectional CIT whereas about 49% showed unidirectional CIT. Of the latter cases, about 30% of the infants were unilateral coordinators, compared with 19% of the adults. In approximately 13% of the interactions, the participants did not appear to engage in statistically significant levels of CIT.

That about 30% fewer adult-infant interactions exhibited bidirectional CIT when the infants were 12 months old is probably explained by the infants' increased interest in and awareness of their environments rather than only the adults in front of them. The reduction no doubt also had to do with the fact that a puppet was part of the interaction.

Interestingly, when the infants were four months old, 41% of the mother-stranger pairs engaged in bidirectional CIT, 41% engaged in unidirectional CIT, and 18% showed no temporal coordination. However, when the infants were 12 months old, only 20% of the mother-stranger interactions showed bidirectional CIT, 52% showed unidirectional CIT, and 28% exhibited no CIT.

A number of other observations about the extent of coordination among the dyads are in order. In general, the strangers engaged the infants in more bidirectional CIT than did the infants' mothers. This may well have been the case because the strangers, who were graduate students trained to interact with infants, felt the need to attend much more closely than the mothers to the infants' actions and vocalizations. Unilaterally, the mothers tended to coordinate more than either the infants or the strangers. The same argument would suggest that after attentively interacting with the infants, the strangers felt much more relaxed and less pressured to coordinate with the mothers who, like themselves, were adult women. The important point, however, is that a sizable proportion of the mothers, strangers and infants engaged in either bidirectional or unidirectional CIT. Equally notable is that the interactions of the mothers and strangers may not represent the degree of coordination achieved by adult pairs in general. They do not, for example, appear to characterize the degree of CIT achieved by the women's interactions in the Crown study (1991) presented earlier.

Special Issues

Autism. Another study, conducted in collaboration with Mary Konstantareas,

---

[11] It is important to note that the percentages of the different types of CIT are based upon statistical significance. For example, the statement that about 9% of the interactions showed no CIT means that the degree of CIT in which they did engage did not reach statistical significance. None of the interactions yielded coefficients that were zero, which would indicate that there was absolutely no CIT. Although my colleagues and I have always used statistical significance as a criterion for judging the occurrence of CIT, it seems to me there is still an open question about the usefulness of such a strategy.

Christopher Webster and Joel Oxman (1982), explored the possibility that the extent of CIT in the interactions of autistic adolescents and young adults is different from that in the interactions of others. Infantile autism has been associated with linguistic, social, cognitive, sensory and perceptual impairments that have been described in some detail elsewhere (*e.g.*, Bartolucci, 1982; Churchill, 1978; Kanner, 1943; Rutter, 1979). Autistic children develop deviant language systems which rarely function to carry consensual meanings. Moreover, Swisher and Hirsch (1972) posited an autonomous temporal organizing system, presumably central to the adequate functioning of the perceptual-sensory system, and it is this system that seems to function inadequately in autistic children. Others (Wing, 1982) have suggested that autistic children have difficulty organizing stimuli sequentially, difficulty that ultimately impairs their social communication. Such difficulty makes problematical the appropriate use of facial, gestural, and vocal expressions in social communication. In addition, the difficulty hinders the temporal integration of an autistic child's behaviors with those of another person. All of this suggests that autistic persons are likely to have relatively low coefficients of CIT. It is possible, however, that an autistic person's increased sensitivity to their difficulty of coordinating with another interacting person's behavior would increase their vigilance, resulting in quite high CIT coefficients.

The 13 autistic individuals who participated in the study ranged in age from 14½ to 20 years. All were from Canada, from either middle or upper socioeconomic backgrounds, and three were women, a gender ratio that is representative of the autistic population (Rimland, 1964). All were attending special programs for autistic persons or regular high school classes. Thus, all of them were relatively articulate and functioned at a higher level than the general autistic population.

As in the adult-child studies described earlier, each of the target subjects interacted with one parent and with the experimenter (JO). After these interactions were completed, the parent and the experimenter conversed. All of the interactions took place in the homes of the autistic persons. Although the pairs were supposed to engage in 20-minute conversations, the difficulties experienced by the autistic subjects made their interactions more like interviews. Indeed, even the parent-experimenter interactions were more like interviews than informal conversations.

In our initial analysis of this study (1982), we estimated degree of CIT by using intra-class correlation coefficients, which is the strategy we used at that time for assessing CIT. However, we have now reanalyzed the data using time-series regression analyses and because they are new, present the coefficients of CIT here (Table 1). Notably, although the averages are not statistically significant, they are consistently

---

Insert Table 1 about here

---

lower than those found in the New York study of adults and infants (Jaffe et al., 1995). For example, the New York study found that the average coefficient of vocalization coordination, over all contexts and conditions, for the adults in that study is .097 and for the 12-month-old infants is .121. The average pause coordination yielded a coefficient of .081 for the adults and .116 for the infants. Switching-pause coordination yielded average coefficients of .091 for the adults and .139 for the infants. Finally, the average ISS and NSS coefficients for the adults and infants are .114 and .118, respectively.

Interestingly, in the original analyses of the study, the parent-experimenter interactions yielded significant intraclass coefficients, suggesting that the pairs engaged in a significant level of CIT. However, Table 1 indicates that the coefficients derived from the parent-experimenter interactions were no higher than those derived from the interactions between either the experimenter or the parents and the autistic subjects. These differences question whether a comparison of session averages, as is done in an intraclass analysis, is an effective way of getting at interpersonal coordination. The CIT coefficients may be consistently lower than the adults and infants in the New York study because these interactions were much more like interviews than like unconstrained conversations. Or the coefficients may, in fact, reflect the problems experienced by the autistic subjects carrying on verbal exchanges and the effects of these problems on the adults in the exchanges. Clearly, the size of the subject sample, the range in the ages of the subjects, the lack of a nonautistic group, and the exploratory nature of the study make it difficult to draw definitive conclusions.

Down's syndrome. While our study of "normal" mother-infant interactions was proceeding in New York, we decided to investigate the extent to which infants with Down's syndrome temporally coordinated their vocal behavior with that of their mothers. We also wanted to examine the relation of CIT to cognitive functioning. The participants in the study, conducted in our Maryland laboratory (Jasnow et al., 1988), were 18 mother-infant pairs. In nine of the pairs, the infants suffered from Down's syndrome (trisomy 21)[12]; the other nine pairs involved nondelayed infants. None of the mothers in the study were clinically depressed or anxious, nor did any of the infants with Down's syndrome present any other relevant health problems. The mother-infant pairs were observed twice, first in the laboratory when the infants were about four, and again when they reached nine months. On each occasion, they engaged in face-to-face play for 12 minutes. The nine-month olds' play included a puppet, and they were given the Bayley MDS. Their interactions were computer coded in terms of the vocal states *speaking turns, pauses, vocalizations, switching pauses*, and *simultaneous speech*. We found that simultaneous speech occurred too infrequently to be further analyzed. The data were subjected to time-series analysis and then to meta analyses.

The results demonstrate that, at four months of age, the infants with Down's syndrome have a lower degree of CIT than that of the nondelayed infants. However, by the time the infants have reached nine months of age, those with Down's syndrome

---

[12]Down's syndrome is a genetically based condition that involves cognitive delay and/or retardation. The average Mental Development Index (of the Bayley MDS) of the infants with Down's syndrome was 64, which is inflated because no score below 50 could be calculated. The average for the nondelayed group was 110.

appear to have caught up with the nondelayed infants in the extent to which they engaged in CIT. Whether or not the MDS is considered a measure of cognition, there is no doubt that the infants with Down's syndrome were cognitively impaired (although to different degrees, of course), and the results suggest that CIT is relatively independent of the level of cognitive functioning. The "catch-up" in CIT from four to nine months is consistent with a good number of other studies about the social behavior of Down's-syndrome infants (Ciccheti & Sroufe, 1976; Serafica & Ciccheti, 1976; Ciccheti & Serafica, 1981; Spiker, 1983). We concluded, from this study, that whatever mechanism is involved in the capacity for CIT appears to be buffered against insults to the organism. Nonetheless, replication of the study with larger groups may provide greater verification.

Predictions of First-Year Outcomes

Among the aims of the New York study described earlier, one was to see whether the extent of CIT exhibited by four-month-old infants could predict their *temperament*, *cognitive functioning,* and the quality of their *attachment* to their mothers at 12 months. Multiple regression equations were used to test these predictions.[13] In addition, because we accounted statistically for all other outcomes, our prediction of each outcome was not dependent upon, or related to, the predictions of the other outcomes.

The analyses revealed significant and complex relationships between the CIT coefficients from the four-month interactions and the three 12-month infant outcomes. Not only do the infants' four-month coefficients predict the 12-month outcomes, but the coefficients of their mothers and the strangers also predict them. Furthermore, the four-month coefficients derived from the mother-stranger conversations predict each of the 12-month outcomes; in some cases, they predict the outcomes more strongly than do the coefficients derived from their interactions with the infants. Here, however, I want to limit my discussion to the prediction of mother-infant attachment.

Mother-infant attachment has to do with the quality of the relationship between them. It was Bowlby (1982), a British psychiatrist, who proposed that the development of mother-infant relationships can be viewed most fruitfully from an ethological, evolutionary perspective. In terms of this perspective, mothers across species engage in certain basic, crucial behaviors that promote or hinder the development, and perhaps survival of their young, including feeding, protection, contact and the like. The perspective is nicely illustrated by much of Harlow's (1958) work with monkeys and his notion of "contact comfort." A pivotal aspect of Bowlby's theory is the construct of a behavioral system (Ainsworth, Blehar, Waters & Wall, 1978) and two critical systems are the attachment system and the exploratory system. The former comes into play whenever the infant feels frightened or threatened. The infant's response to these feelings is ordinarily to seek the close contact and protection with the mother (or

[13]Details about the order of the dependent variables and covariates and independent variables in each of the equations is presented in Jaffe et al. (1995).

caretaker). The exploratory system, which is activated once the threat is passed, allows the infant to explore his environment. Ainsworth devised a procedure – called the Strange Situation – for assessing the type and quality of mother-infant attachment. The procedure involves the infant in eight different episodes, the most important being the two in which the infant and mother are reunited. On the basis of their behavior during the reunion episodes, Ainsworth proposed three categories of mother-infant attachment, each with a number of subcategories. Briefly, categories *A* and *C* characterize two types of insecure attachment; category *B* depicts secure attachment. Anxious, avoidant infants are classified as *A* babies because of their minimal crying and avoidance of their mothers during the reunion episodes; they engage in relatively low proximity seeking and contact maintenance. The infants in the *C* category appear to be very distressed by the previous separation and tenaciously seek close contact with their mothers upon reunion. At the same time, however, they exhibit angry resistance to their mothers, and cannot be comforted. Those infant classified as secure (*B*), while also distressed by the prior separation, seek moderate proximity and contact with their mothers and can be comforted by them. Mary Main (1983) added a *D* category, consisting of infants who are considered disorganized and disoriented, and characterized by a constant shifting of attachment strategies in the face of threat or fear.

We attempted to predict not only the four attachment categories, *A, B, C,* and *D*, but also a continuous estimate of attachment derived by discriminant function analysis by Richters, Waters, and Vaughn (1988) from the behavioral scales used to code the Strange Situation. The results of our analyses indicate that we *can* predict mother-infant attachment at 12 months from their four-month coefficients of CIT. In fact, we can predict all four categories of attachment, but we need to analyze both the mother-infant and stranger-infant interactions to predict the full range of attachment variations. My colleagues Jasnow and Jaffe (personal communications, August, 1994) have conjectured that the strangers presented a challenge to the infants that elicited more specific demonstrations of attachment behavior. Jaffe (personal communication, August, 1995) is in the process of formulating a mathematical expression to describe what appears to be a lawful relationship between degree of challenge and the coefficients of CIT.

The most general finding is that a *high* degree of coordinated timing bodes ill for the infant. Mother-infant interaction, both in the home and laboratory, presents a consistent and bidirectional picture in which the highest and mutual tracking of the mother and infant occurred among the pairs classified as *D*, or disoriented and/or disorganized relationships. The next most general finding is that *midrange* timing predicts relatively secure relationships and the low and high ranges predict the insecure and so-called disorganized attachment relationships.

It has been argued by Mary Main (1983) and, particularly, by Mary Sue Moore (personal communication, 1994), that the mothers in the *D* relationships are highly unpredictable in their responses to their infants. Moore suggests, on the basis of extended and careful observations of such relationships, that a *D* mother may respond positively to her infant at one moment but be momentarily dissociated at the next instant or look at the infant with a fleeting, but distinct expression of hate and/or rage. That expression may last for only an instant but is intercepted by the infant. Faced with sudden and inexplicable changes in the mother's responses, the infant's shifting strategies and/or apparent disorientation may represent an adaptive tactic. One hypothesis is that the mother is herself struggling with an interactive history of abuse and/or insecurity. In the light of the behaviors in which these mothers and infants engage, the high tracking or coordinated timing may well be a compensatory effort on the parts of the mothers and infants to make the interaction more manageable.

Apart from those relationships considered disorganized, it seems as if the insecure relationships result in more or less than optimal coordination. Perhaps the greater coordination results from the infants' feeling, (with their mothers or, especially, with the strangers), that more vigilant monitoring is needed to meet the challenge of an unknown person and/or a new context (the laboratory). Conversely, the lesser coordination may represent an effort to avoid interpersonal contact.

Some Current and Possible Directions

Other colleagues and I intend to assess the extent of CIT in the interactions of 18-month olds with their mothers in play sessions and the degree to which the coefficients of CIT appear to have been influenced by, or to fluctuate with the attachment relationship. The investigation will also allow us to follow the developmental course of CIT, which we now believe decreases in magnitude from four to 12 months.

The question of whether CIT is a neurobiological mechanism is still viable. I had suggested earlier that it would be useful to approach it with relevant animal studies. One such study might involve the use of squirrel monkeys, the only type of monkey that chatters in pairs, rather than in small groups. A number of investigators have, in fact, studied the relation between the monkeys' temporal patterning of their vocal interactions and their social behavior. One investigation (Biben, Symmes, & Masataka, 1986) found that the temporal organization of their interactions were associated with their affiliative behaviors. A later study (Masataka & Biben, 1987) found that, specifically, the latency durations of their interactions describe affiliative exchanges but not other types of exchanges. The authors assert that it is one of the temporal rules that characterize the affiliative exchanges of squirrel monkeys and serves a regulatory function in the exchange. The intriguing aspect of this finding is that it suggests that for squirrel monkeys, as for humans, the chronometry of their interactions may be related to the nature of their relationships.

Crown and her student and I (Crown, Wolf, & Feldstein, 1990) wondered whether that particular vocal state (latencies) served a similar function in human interactions. We consider the silence between the end of one turn and the start of another a *switching pause*, which, unlike a latency, is assigned to the individual whose turn has ended. In this study, we reassigned the switching pauses of the interactions of a group of 76 students in order to convert them to latencies. We then compared them to the students' score on the affection scales of the FIRO-B Interpersonal Relations Orientation Behavior

(Schutz, 1958). The two affection scales were used to estimate need for affiliation. The comparison yielded a significant positive relationship between the latencies and the affection scores. Although the finding needs to be replicated, it appears as though latencies may serve a similar function for humans and squirrel monkeys. However, whereas shorter latencies are associated with more affiliative behaviors in monkeys, the opposite is true for humans. Those students with a greater need for affection/affiliation had longer latencies.

If very high and very low levels of CIT in the interactions of four-month-old infants signal the development of less than optimal attachment relationships, they may also signal the development of problematic social relationships in general. A study that has yet to be done is to follow the infants until they have been in school a few years and examine the nature of their social relationships with other children. It might also be useful to observe, at that time, the extent to which they engage in CIT with their peers and to compare the findings with their four-month coefficients of CIT.

Relatively early in our research, Jaffe initiated a search for mathematical models that would usefully describe the data (Jaffe, Cassotta, & Feldstein, 1964; Jaffe, Feldstein, & Cassotta, 1967a; Jaffe, Feldstein, & Cassotta, 1967b). At the same time, I wondered whether we could devise a recursive model (Feldstein, 1972) to describe the cluster of vocal states, but found that the mathematical and recursive models we had formulated were essentially identical transformations. We viewed the vocal states as elements within a system. Pauses, vocalizations, switching pauses, and simultaneous speech are, after all, components of a single dialogue, a single dyadic system. They are conceptually related even though several of them may be empirically uncorrelated. He proposed the six-state independent decision model (Jaffe & Feldstein, 1970), which states that the *independent* decision to do something (in terms of vocal behavior) that each participant makes at any moment is *dependent* upon the *dyadic* state of the previous moment. Another important and related model to describe mother-infant interaction was proposed by Thomas and Malone (1979). Crown and Flaspohler conducted a set of studies (Crown, Flaspohler, & Jaffe, 1993; Crown, Flaspohler, Feldstein, Jaffe, & Beebe, 1995) that compared the two versions of the original Markovian models and then compared the conditional dependence model with that proposed by Thomas and Malone. The studies also compared the models across the ages of 6 weeks, 4 months, and 12 months, using the infants' visual behavior when they were six weeks old. They found that the conditional dependence model fit less well as the age of the infants increased. Interpretation of the findings associated with the Thomas and Malone model was confusing because, although the model presumably separates the effects of the speaker from the influence of the listener ("self" and "other" effects), it does so more in name than in fact. In any case, there are some interesting questions about the conditional dependence model and other possible mathematical models that have yet to be answered. If, for example, the models fit less well as age increases, is that because more hidden states come into play as the interactive capabilities of speakers mature?

Concluding Remarks

My purpose in this chapter was to provide a simple description of what we mean by coordinated interpersonal timing, how we assess it, and the ways in which we have recently investigated it. I also have tried to show that the sequences of sounds and silence that represent our streams of speech carry a good bit of information. But it is not semantic information and hardly any syntactic information (Gerstman, Feldstein, & Jaffe, 1967). Instead, it carries information about certain aspects of an individual's personality (Feldstein, Alberti, BenDebba, & Welkowitz, 1974), what kinds of interpersonal relationships pairs of individuals have, and what the participants in each pair feel about each other. It seems particularly true that in many casual conversations the words that are used are less important than the ways in which they are said and the temporal structure in which they are embedded. It is also true that a chronometric analysis describes only one facet, or channel, of the full "package" of speech. Other important facets, such as loudness, pitch, syntactics and semantics (and perhaps the gestures associated with speech) are being studied by others. Another approach that has been taken (e.g., Siegman & Feldstein, 1985 attempts to integrate the channels to provide a picture of the total communicative act. It is a very worthwhile approach, but to accomplish it with the precision and objectivity that it merits would require a range of sophisticated technological equipment and a group of scientists from diverse disciplines. It is likely that such a combination is not yet feasible although it is certainly worth considering. Nevertheless, apart from this approach, it remains important to clarify the contributions of the separate channels to the "package."

Finally, an assumption underlying the use of an automated procedure for examining the chronometric channel of speech is that objectivity and precision are attributes that characterize a scientific endeavor. The move from n^{th} order conjectures about the meanings of behavior to the actual observation of behavior represented a significant advance in scientific psychology. The move from observing behavior to objectively and precisely measuring it represents another advance; it becomes possible to *investigate* the meanings of behavior.

References

Ainsworth, M. D. S., Blehar, M. C., Waters, e., & Wall, s. (1978). *Patterns of attachment: A psychological study of the Strange Situation.* Hillsdale, NJ: Erlbaum Associates.

Allen, R. R., Anderson, S. & Hough, J. (1968). *Speech in American society.* Columbus, OH: Charles E. Merrill

Allen, T., Walker, K., Symonds, L. & Marcell, M. (1977). Intrasensory and intersensory perception of temporal sequences during infancy. *Developmental Psychology, 13*(3), 225-229.

Anderson, S. W., Jaffe, J. & Chen, T. (1992). Speech timing: Patterns and prognosis in mood disorders. *Abstracts from the 1992 Annual Meeting of the AAAS*, p. 142.

Bartolucci, G. (1982). Formal aspects of language in childhood autism. *Advances in Childhood Behavioral Analysis and Therapy, 2,* 159-185.

Berscheid, E. (1985). Interpersonal attraction. In G. Lindzey and E. Aronsmith (Eds.), The handbook of social psychology. New York: Random House.

Biben, M., Symmes, D., Masataka, N. (1986). Temporal and structural analysis of affiliative vocal exchanges in squirrel monkeys (Saimiri, sciureus). *Behavior, 49,* 227-267.

Bowlby, J. (1982). *Attachment* (2nd ed.).New York: Basic Books.

Bryant, P. E., Jones, P., Claxton, B., & Perkins, G. M. (1972). Recognition of shapes across modalities by infants. *Nature, 240,* 303-304.

Byers, P. (1976). Biological rhythms as information channels in interpersonal communication behavior. In P. P. G. Bateson & P. H. Klopfer (Eds.), *Perspectives in ethology.* (Vol 11, pp. 135-164). New York: Plenum.

Cappella, J. N. & Greene, J. O.(1982). A discrepancy-arousal explanation of mutual influencein expressive behavior for adult and infant-adult interaction. *Communication Monographs,* 19, 89-114.

Cassotta, L., Feldstein, S., & Jaffe, J. (1964). AVTA: A device for automatic vocal transaction analysis. *Journal of the Experimental Analysis of Behavior, 7,* 99-104.

Cassotta, L., Feldstein, S., & Jaffe, J. (April, 1967b). *Stability, influence and congruence in dyadic vocal interaction.* Paper read at the Eastern Psychological Association, Boston.

Chapple, E. D. (1970). *Culture and biological man: Explorations in behavioral anthropology.* New York: Holt, Rinehart and Winston.

Churchill, D. W. (1978). *Language of autistic children.* New York: John Wiley.

Ciccheti, D., and Serafica, F. (1981). Interplay among behavioral systems: Illustrations for the study of attachment, affiliation and wariness in young children with Down's syndrome. *Developmental Psychology, 17,* 36-49.

Ciccheti, D., and Sroufe, A. (1976). The relationship between cognitive development in Down's syndrome infants. *Child Development, 47*: 920-929.

Cohn, J. F. & Tronick, E. Z. (1988). Mother-infant face-to-face interaction: Influence is bidirectional and unrelated to periodic cycles in either partner's behavior. *Developmental Psychology, 24,* 386-392.

Crown, C. L. (1982). Impression formation and the chronology of dyadic interactions. In M. Davis (Ed.), *Interaction rhythms: Periodicity in communicative behavior.* New York: Human Sciences.

Crown, C. L. (1991). Coordinated interpersonal timing of vision and voice as a function of interpersonal attraction. *Journal of Language and Social Psychology, 10,* 29-45.

Crown, C. L., & Feldstein, S. (1985). Psychological correlates of silence and sound in conversational interaction. In D. Tannen & M. Saville-Troike (Eds.), *Perspectives on silence*. Norwood, NJ: Ablex, 31-54.

Crown, C. L. & Feldstein, S. (1991). The perception of speech rate from the sound-silence patterns of monologues. *Journal of Psycholinguistic Research, 20,* 47-63.

Crown, C. L., Feldstein, S., & Bond, R. N. (1982). Effects of simultaneous speech on interpersonal perception. Paper read at the Eastern Psychological Association, Baltimore.

Crown, C. L., Flaspohler, D. C., & Jaffe, J. (May, 1993). A Markovian analysis of mother-infant play at 6 weeks, 4 months and 12 months. Poster presented at the International Conference on Infant Studies, New Orleans.

Crown, C. L., Wolf, M. L., & Feldstein, S. (March, 1990). *Affiliation and interaction chronography in humans and squirrel monkeys*. Poster presented at the Eastern Psychological Association, Philadelphia.

Crown, C. L., Flaspohler, D. C., Feldstein, S. Jaffe, J., Beebe, B. & Jasnow, M. D. (1995). *Mathematical models for coordinated interpersonal timing in mother-infant interactions in the first year of life*. Manuscript submitted for publication.

Deaux, K. (1977). Sex differences. In T. Blass (Ed.), *Personality variables in social behavior* (pp. 357-377). Hillsdale, NJ: Erlbaum Associates.

Feldstein, S. (1972). Temporal patterns of dialogue: Basic research and reconsiderations. In A. W. Siegman & B. Pope (Eds.), *Studies in dyadic communication*. New York: Pergamon Press, 91-113.

Feldstein, S. (1976). Rate estimates of sound-silence sequences in speech. *Journal of the Acoustical Society of America, 60* (Suppl. No.1) S46. (Abstract)

Feldstein, S. (1982). Impression formation in dyads: The temporal dimension. In M. Davis (Ed.), *Interaction rhythms: Periodicity in communicative behavior*. New York: Human Sciences, 207-224.

Feldstein, S. (April, 1991). Vocal behavior in the interactions of preadolescent friends and acquaintances. In T. Field (Chair), *Interaction rhythms across development*. Symposium presented at the Society for Research in Child Development, Seattle, WA.

Feldstein, S., Alberti, L., BenDebba, M., & Welkowitz, J. (September, 1974). *Personality and simultaneous speech*. Paper read at the American Psychological Association, New Orleans.

Feldstein, S., BenDebba, M., & Alberti, L. (1974). Distributional characteristics of simultaneous speech in conversation. *Journal of the Acoustical Society of America, 55* (Suppl.) S42. (Abstract)

Feldstein, S., Crown, C. L., & Jaffe, J. (1991). Expectation and extraversion: Influencing the perceived rate of tone-silence sequences. *Bulletin of the Psychonomic Society, 29,* 395-398.

Feldstein, S., Jaffe, J., & Cassotta, L. (1967). The effect of mutual visual access upon conversational time patterns. *American Psychologist, 23,* 595. (Abstract)

Feldstein, S., Konstantareas, M., Webster, C. D., & Oxman, J. (1982). The chronography of interactions with autistic speakers: An initial report. *Journal of Communication Disorders, 15,* 451-460.

Feldstein, S., Siegman, A. W., Simpson, S., Barkley, S., & Kobren, R. (April, 1984). *Assessing coronary prone behavior from the temporal structure of speech*. Paper read at the Eastern Psychological Association, Baltimore.

Feldstein, S., & Sloan, B. (1984). Actual and stereotyped speech patterns of extraverts and introverts. *Journal of Personality, 52,* 188-204.

Feldstein, S. & Weingartner, H. (1981). Speech and psychopharmacology. In J. K. Darby, Jr.

(Ed.), *Speech evaluation in psychiatry (pp.369-396)*. New York: Grune & Stratton.

Feldstein, S., & Welkowitz, J. (1978). A chronography of conversation: In defense of an objective approach. In A. W. Siegman & S. Feldstein (Eds.), *Nonverbal behavior and communication* (pp.329-378). Hillsdale, NJ: Erlbaum Associates. Second edition, 1987, pp. 435-499.

Freides, D. (1974). Human information processing and sensory modality: Cross-modal functions, information complexity, memory, and deficit. *Psychological Bulletin*, 81, 284- 310.

Garvey, C., & BenDebba, M. (1974). Effects of age, sex and partner on children's dyadic speech. *Child Development, 45*, 1159-1161.

Gerstman, L. J., Feldstein, S., & Jaffe, J. (November, 1967). *Syntactical versus temporal cues for speaker switching in natural dialogue*. Paper read at the Acoustical Society of America, Florida.

Giles H., Taylor, D. M., & Bourhis, R. Y. (1973). Towards a theory of interpersonal accommodation through language: Some Canadian data. *Language in Society, 2*, 177-192.

Glaister, J., Feldstein, S. & Pollack, H. (1980). Chronographic speech patterns of acutely psychotic patients: A preliminary note. *Journal of Nervous and Mental Disease*, *168*, 219-223.

Greden, J. F. & Carroll, B. J. (1980). Decrease in speech pause time with treatment of endogenous depression. *Biological Psychiatry*, *15*, 575-587.

Harlow, H. F. (1958). The nature of love. *American Psychologist*, 13, 673-685.

Homans, G. C. (1950). *The human group*. New York: Harcourt Brace.

Jaffe, J., & Feldstein, S. (1970). *Rhythms of dialogue*. Academic.

Jaffe, J., Cassotta, L., & Feldstein, S. (1964). A Markovian model of time patterns in speech. *Science*, *144*, 884-886.

Jaffe, J., Feldstein, S., & Cassotta, L. (1967a). Markovian models of dialogic time patterns. *Nature*, *216*, 93-94.

Jaffe, J., Feldstein, S., & Cassotta, L. (1967b). A stochastic model of speaker switching in natural dialogue. In K. Salzinger & S. Salzinger (Eds.), *Research in verbal behavior and some neurophysiological implications*. New York: Academic, 281-294.

Jaffe, J., Feldstein, S., Beebe, B., Crown, C. L., & Jasnow, M. D. (1995). *Coordinated interpersonal timing and infant social behavior.* Manuscript in preparation.

Jasnow, M. D., Crown, C. L., Feldstein, S., Taylor, L., Beebe, B. & Jaffe, J. (1988). Coordinated interpersonal timing of Down-syndrome and nondelayed infants with their mothers: Evidence for a buffered mechanism of social interaction. *Biological Bulletin*, *175*, 355-360.

Jones, B. & Robinson, T. (1973). Sensory integration in normal and retarded children. *Developmental Psychology*, *9*, 178-183.

Kanner, L. (1943). Autistic disturbance of affective contact. *The Nervous Child*, *2*, 217-250.

Kirkpatrick B., Carter, C. S., Newman, S. W., & Insel, T. R. (1994). Axon-sparing lesions of the medial nucleus of the amygdala decrease affiliative behaviors in the prairie vole (*Microtus orchraster*): Behavioral and anatomical specificity. *Behavioral Neuroscience*, *108*, 501-513.

Krüger, H-P., Rausche, A., Rimkus, U., & Voorath, M. (1991). The Logoport: Handbook. Unpublished manuscript, Department of Psychology, University of Würzburg, Germany.

LaFrance, M. (1979). Nonverbal synchrony and rapport: Analysis by the cross-lag technique. *Social Psychology Quarterly, 42,* 66-70.

Lewkowicz, D. J. (1989). The role of temporal factors in infant behavior and development. In I. Levin & D. Zakay (Eds.), *Time and human cognition*. North-Holland; Elsevier Science Publishers, 9-62.

Marcus, E. S., Welkowitz, J., Feldstein, S., & Jaffe, J. (1970). Psychological differentiation and the congruence of temporal speech patterns. Paper presented at the meeting of the Eastern Psychological Association, Atlantic City.

Martin, J. A. (1981). A longitudinal study of the consequences of early mother-infant interactions: A microanalytic approach. *Monographs of the Society for Research in Child Development, 46* (3, Serial No. 190).

Martindale, D. A. (1971). *Effects of environmental context in negotiating situations: Territorial dominance behavior in dyadic interactions.* Unpublished doctoral dissertation, City University of New York.

Masataka, N., & Biben, M. (1987). Temporal rules regulating affiliative vocal exchanges of squirrel monkeys. *Behavior, 101,* 311-319.

Matarazzo, J. D. & Wiens, A. N. (1967) Interviewer influence on durations of interviewee silence. *Journal of Experimental Research in Personality, 2,* 56-69

Miller, J. L. (1987). Rate-dependent processing in speech perception. In A. e. Ellis (Ed.), *Progress in the psychology of language*: Vol 3 (pp. 119-157). Hillsdale, N.J.: Erlbaum.

Miller, N., Maruyama, G., Beaber, R. J., & Valone, K. (1976). Speed of speech and persuasion. *Journal of Personality and Social Psychology, 34,* 615-624.

Natale, M. (1975). Convergence of mean vocal intensity in dyadic communication as a function of social desirability. *Journal of Personality and Social Psychology, 32,* 790-804.

Natale, M., Enton, E., & Jaffe, J. (1979). Vocal interruptions in dyadic communication as a function of speech and social anxiety. *Journal of Personality and Social Psychology, 37,* 865-878.

Newcomb, T. M. (1953). An approach to the study of communicative acts. *Psychological Review, 15,* 393-404.

Ramsay, R. W. (1966). Personality and speech. *Journal of Personality and Social Psychology, 4,* 936-945.

Ramsay, R. W. (1968). Speech patterns and personality. *Language and Speech, 11,* 54-63.

Richters, J. E., Waters, E. & Vaughn, B. E. (1988). Empirical classification of infant-mother relationships from interactive behavior and crying during reunion. *Child Development, 59,* 512-522.

Rimland, B. (1964). *Infantile autism.* New York: Appleton-Century-Crofts.

Rose, S. A., Gottfried, A. W. , & Bridger, W. H. (1978) Cross-modal transfer in infants: Relationship to prematurity and socioeconomic background. *Developmental Psychology, 14,* 643-652.

Rutter, M. (1979). Language, cognition, and autism. In R. Kathman (Ed.), *Congenital and acquired cognitive disorders.* The Association for Research in Nervous and Mental Disease, 57, 247-264.

Schaffer, H. R. (1977). *Studies in mother-infant interaction.* New York: Academic.

Scherer, K. R. (1979). Personality markers in speech. In K. R. Scherer & Giles (Eds.), *Social markers in speech*. Cambridge: Cambridge University Press.

Schutz, W. (1958). *FIRO: A three dimensional theory of interpersonal behavior*. New York: Rinehart and Co., Inc.

Schwab, J. J. (1960). What do scientists do? *Behavioral Science, 5*, 1-27.

Serafica, F., & Ciccheti, D. (1976). Down's syndrome children in a strange situation: Attachment and exploration behaviors. *Merrill-Palmer Quarterly, 22*: 137-150.

Siegman, A. W. (1987). The telltale voice. In A. W. Siegman & S. Feldstein (Eds.), *Nonverbal behavior and communication*. Hillsdale, NJ: Erlbaum Associates, 329-378. Second edition, 1987, pp. 351-420.

Siegman, A. W. & Feldstein, S. (Eds.). (1985). *Multichannel integrations of nonverbal behavior*. Hillsdale, NJ: Erlbaum Associates.

Siegman, A. W., Feldstein, S., Tommaso, C. T., Ringel, N. & Lating, J. 87). Expressive vocal behavior and the severity of coronary artery disease. *Psychosomatic .ledicine, 49*, 545-561.

Spiker, D. 1983. Early intervention for young children with Down's syndrome: New directions for enhancing parent child synchrony. In S. Peuschel & J. Rynders (Eds), *Down's syndrome: Advances in biomedicine and the behavioral sciences*. Cambridge, MA: Ware Press.

Swisher, L. & Hirsch, I. (1972). Brain damage and the ordering of temporally successive stimuli. *Neuropsychologia, 10*, 137-152.

Teasdale, J. D., Fogarty, S. J., & Williams, J. M. G. (1980). Speech rate as a measure of short-term variation in depression. *British Journal of Social and Clinical Psychology, 19*, 271-278.

Turkewitz, G. & McGuire, I. (1978). Intersensory functioning during early development. *International Journal of Mental Health*, 7, 165-182.

Warner, R. M. (1988). Rhythm in social interactions. In J. E. McGrath (Ed.), *The social psychology of time: New perspectives* (pp. 63-87). Beverly Hills: Sage.

Warner, R. M. (1992). Sequential analysis of social interaction: Assessing internal versus social determinants of behavior. *Journal of Personality and Social Psychology, 63*, 51-60.

Webb, J. T. (1972). Interview synchrony: An investigation of two speech rate measures. In A. W. Siegman & B. Pope (Eds.), *Studies in dyadic communication* (pp. 115-133.). New York: Pergamon.

Welkowitz, J., & Feldstein, S. (1970). Relation of experimentally manipulated interpersonal perception and psychological differentiation to the temporal patterning of conversation. *Proceedings of the 78th Annual Convention of the American Psychological Association, 5*, 387-388.

Welkowitz, J., Bond, R. N., & Feldstein, S. (1984). Conversational time patterns of Hawaiian children as a function of ethnicity and gender. *Language and Speech, 27*, 173-191.

Welkowitz, J., Bond, R. N., & Feldstein, S. (1984). Gender and conversational time patterns of Japanese-American adults and children in same- and mixed-gender dyads. *Journal of Language and Social Psychology, 3*, 127-138.

Welkowitz, J., Cariffe, G., & Feldstein, S. (1976). Conversational congruence as a criterion of socialization in children. *Child Development, 47*, 269-272.

Welkowitz, J., Bond, R. N., Feldman, L., & Tota, M. E. (1990). Conversational time patterns and mutual influence in parent-child interactions: A time series approach. *Journal of Psycholinguistic Research, 19,* 221-243.

Welkowitz, J., Feldstein, S., Finkelstein, M., & Aylesworth, L. (1972). Changes in vocal intensity as a function of interspeaker influence. *Perceptual and Motor Skills, 35,* 715-718.

Wing, L. (1982). *The handicaps of autistic children: An aid to diagnosis*. London: National Autistic Society.

Table 1

Descriptive Statistics of the Coefficients of Coordinated Interpersonal Timing (r^2s) from the Interactions of Autistic Adolescents and Young Adults with a Parent and with an Experimenter

Participant		T	V	P	SP	ISS	NSS
Autistic Person	***M***	.135	.067	.062	.053	.041	.066
	SD	.127	.030	.022	.026	.014	.053
with							
	M	.063	.062	.071	.063	.065	.073
Experimenter	***SD***	.041	.037	.033	.024	.028	.029
Autistic Person	***M***	.084	.085	.071	.056	.064	.055
	SD	.094	.026	.027	.046	.027	.030
with							
	M	.092	.072	.063	.093	.065	.082
Parent	***SD***	.077	.028	.029	.052	.028.	.084
Parent	***M***	.063	.065	.079	.068	.095	.073
	SD	.051	.027	.034	.048	.034	.021
with							
	M	.095	.070	.059	.051	.055	.063
Experimenter	***SD***	.125	.041	.024	.043	.027	.024

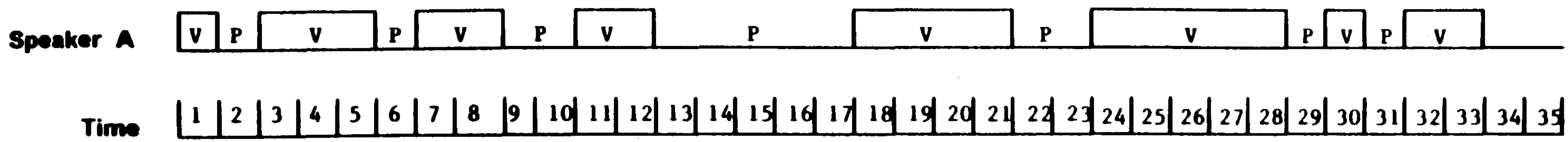

Fig. 1. A diagramatic segment of a monologue. The time line is in units of one quarter second. Along Participant *A*'s line, the closed boxes represent sounds, or vocalizations (V). The unclosed spaces between them are silences, or pauses (P). Note that the silence above time units 34 and 35 cannot be classified because it is not known from the speech segment whether or not Speaker *A* will continue to speak.

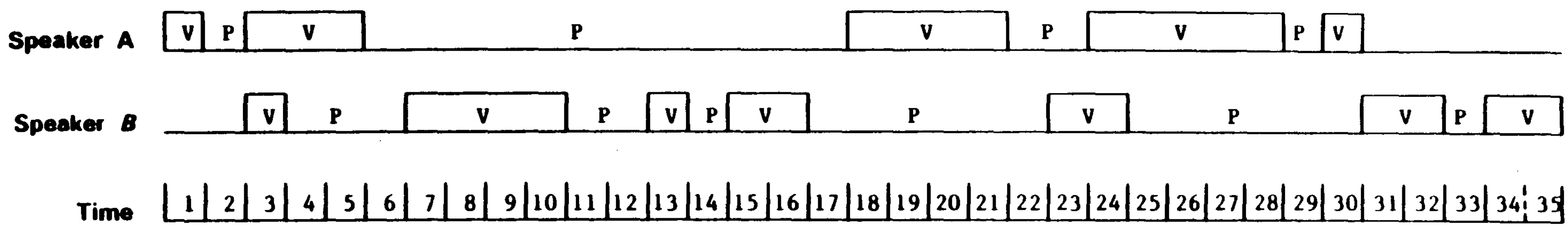

Fig. 2. A diagrammatic segment of two persons talking. The time line at the bottom is divided into 250-msec units. **P** stands for *pauses*, and **V** for *vocalizations*.

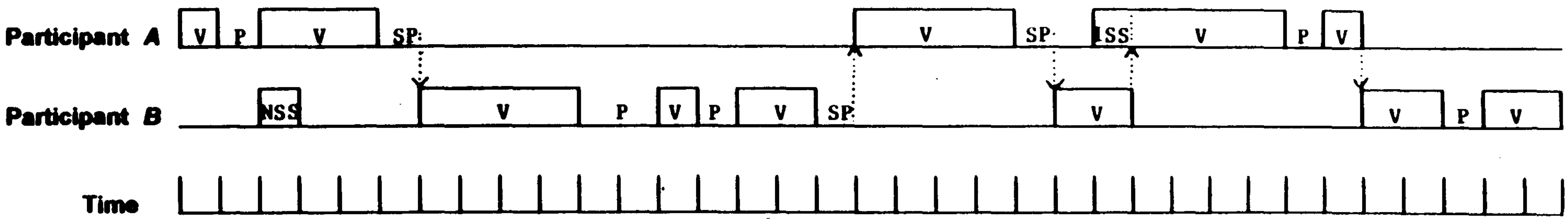

Fig. 3. A diagrammatic representation of a conversational segment. The time axis is divided into 250-msec. units. V = vocalization, P = pause, SP = switching pause, NSS = noninterruptive simultaneous speech, and ISS = interruptive simultaneous speech. The arrows that point down denote the end of Participant A's turns, those that point up denote the end of Participant B's turns.